PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/887,977*

DATE: 03/09/98
TIME: 14:52:15

INPUT SET: S24076.raw

SECTOR

> This Raw Listing contains the General
> Information Section and up to the first 5 pages.

```
 1                          SEQUENCE LISTING
 2
 3  (1)     General Information:
 4
 5       (i) APPLICANT: Wang, Wei
 6                      Gish, Kurt C.
 7                      Schall, Thomas J.
 8                      Vicari, Alain P.
 9                      Zlotnik, Albert
10
11      (ii) TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
12
13     (iii) NUMBER OF SEQUENCES: 19
14
15      (iv) CORRESPONDENCE ADDRESS:
16           (A) ADDRESSEE: DNAX Research Institute
17           (B) STREET: 901 California Avenue
18           (C) CITY: Palo Alto
19           (D) STATE: California
20           (E) COUNTRY: USA
21           (F) ZIP: 94304-1104
22
23       (v) COMPUTER READABLE FORM:
24           (A) MEDIUM TYPE: Floppy disk
25           (B) COMPUTER: IBM PC compatible
26           (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27           (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
28
29      (vi) CURRENT APPLICATION DATA:
30           (A) APPLICATION NUMBER: US 08/887,977
31           (B) FILING DATE: 03-JUL-1997
32           (C) CLASSIFICATION:
33
34     (vii) PRIOR APPLICATION DATA:
35           (A) APPLICATION NUMBER: US 60/021,644
36           (B) FILING DATE: 05-JUL-1996
37
38     (vii) PRIOR APPLICATION DATA:
39           (A) APPLICATION NUMBER: US 60/028,329
40           (B) FILING DATE: 11-OCT-1996
41
42    (viii) ATTORNEY/AGENT INFORMATION:
43           (A) NAME: Ching, Edwin P.
44           (B) REGISTRATION NUMBER: 34,090
45           (C) REFERENCE/DOCKET NUMBER: DX0589K1
46
```

ENTERED

INPUT SET: S24076.raw

```
47          (ix) TELECOMMUNICATION INFORMATION:
48               (A) TELEPHONE: 650-852-9192
49               (B) TELEFAX: 650-496-1200
50
51
52     (2) INFORMATION FOR SEQ ID NO:1:
53
54          (i) SEQUENCE CHARACTERISTICS:
55               (A) LENGTH: 1034 base pairs
56               (B) TYPE: nucleic acid
57               (C) STRANDEDNESS: single
58               (D) TOPOLOGY: linear
59
60          (ii) MOLECULE TYPE: cDNA
61
62
63          (ix) FEATURE:
64               (A) NAME/KEY: CDS
65               (B) LOCATION: 94..525
66
67
68          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
69
70     AGGCTACAAG CAGGCACCAG CTCTCAGGAC CAGAAAGGCA TTGGTGGCCC CCTTAAACCT          60
71
72     TCAGGTATCT GGAGAGGAGA TCTAACCTTC ACT ATG AAA CTG TGG CTT TTT GCC         114
73                                        Met Lys Leu Trp Leu Phe Ala
74                                         1               5
75
76     TGC CTG GTT GCC TGT TTT GTT GGG GCC TGG ATG CCG GTT GTC CAT GCC         162
77     Cys Leu Val Ala Cys Phe Val Gly Ala Trp Met Pro Val Val His Ala
78          10                  15                  20
79
80     CAA GGT GCC TTT GAA GAC TGC TGC CTG GGT TAC CAG CAC AGG ATC AAA         210
81     Gln Gly Ala Phe Glu Asp Cys Cys Leu Gly Tyr Gln His Arg Ile Lys
82          25                  30                  35
83
84     TGG AAT GTT CTC CGG CAT GCT AGG AAT TAT CAC CAG CAG GAA GTG AGT         258
85     Trp Asn Val Leu Arg His Ala Arg Asn Tyr His Gln Gln Glu Val Ser
86     40                  45                  50                  55
87
88     GGA AGC TGC AAC CTA CGT GCT GTG AGA TTC TAC TTC CGC CAG AAA GTA         306
89     Gly Ser Cys Asn Leu Arg Ala Val Arg Phe Tyr Phe Arg Gln Lys Val
90                  60                  65                  70
91
92     GTG TGT GGG AAT CCA GAG GAC ATG AAT GTG AAG AGG GCG ATA AGA ATC         354
93     Val Cys Gly Asn Pro Glu Asp Met Asn Val Lys Arg Ala Ile Arg Ile
94                  75                  80                  85
95
96     TTG ACA GCT AGG AAA AGG CTA GTC CAC TGG AAG AGC GCC TCA GAC TCT         402
97     Leu Thr Ala Arg Lys Arg Leu Val His Trp Lys Ser Ala Ser Asp Ser
98               90                  95                  100
99
```

INPUT SET: S24076.raw

```
100   CAG ACT GAA AGG AAG AAG TCA AAC CAT ATG AAG TCC AAG GTG GAG AAC        450
101   Gln Thr Glu Arg Lys Lys Ser Asn His Met Lys Ser Lys Val Glu Asn
102       105                 110                 115
103
104   CCC AAC AGT ACA AGC GTG AGG AGT GCC ACC CTA GGT CAT CCC AGG ATG        498
105   Pro Asn Ser Thr Ser Val Arg Ser Ala Thr Leu Gly His Pro Arg Met
106   120                 125                 130                 135
107
108   GTG ATG ATG CCC AGA AAG ACC AAC AAT TAAGTTAATT ACTCAGAGTA              545
109   Val Met Met Pro Arg Lys Thr Asn Asn
110                   140
111
112   AGCACCAGCT GGAGGATGGG CGGAGTCTGC TGAAGTGCTG TCTTCTAGGC ATGCCAGTGC      605
113
114   CAATGAACTC ACTGAAGCTA CAGTTTCCTG TACAAGACCA GACCCACCAA CGTCTCAGCA      665
115
116   TGTACGAGGA AGGAACTACT GCGCTAAAGG CCCTCCCACT CACCAAGGAG CTATTGGCTA      725
117
118   TTGATGATTG CTGAGGGAAG GGAGTAATTT TTTTTCTCTT TCTGAAGTGT GACTTGAGTA      785
119
120   AATTGCCCAT AGTTCAGTAT ATAATCCCCA ACCTGTGCTC AGGCAAGCAA CCCTAATTAA      845
121
122   ATGCAATAGC CACATACAAA AGAAGAGGAT ATGAATAGTT TGGTAGGAGG GGCTTGTTAG      905
123
124   GAAGAAGACA TTAACAGGAG AGAGAGGAGC GAGAGGATAG TGAGTGTGTG AGAGTGCCTG      965
125
126   CACGTGTGAA ATGGTCAAAG AATTAAAAAA TAAAAACTTA AAAAGCTATT AAAAAGTAAA      1025
127
128   AAAAATAAA                                                              1034
129
130
131   (2) INFORMATION FOR SEQ ID NO:2:
132
133           (i) SEQUENCE CHARACTERISTICS:
134               (A) LENGTH: 144 amino acids
135               (B) TYPE: amino acid
136               (D) TOPOLOGY: linear
137
138       (ii) MOLECULE TYPE: protein
139
140       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
141
142   Met Lys Leu Trp Leu Phe Ala Cys Leu Val Ala Cys Phe Val Gly Ala
143     1               5                  10                  15
144
145   Trp Met Pro Val Val His Ala Gln Gly Ala Phe Glu Asp Cys Cys Leu
146               20                  25                  30
147
148   Gly Tyr Gln His Arg Ile Lys Trp Asn Val Leu Arg His Ala Arg Asn
149           35                  40                  45
150
151   Tyr His Gln Gln Glu Val Ser Gly Ser Cys Asn Leu Arg Ala Val Arg
152           50                  55                  60
```

INPUT SET: S24076.raw

```
153
154  Phe Tyr Phe Arg Gln Lys Val Val Cys Gly Asn Pro Glu Asp Met Asn
155   65              70              75              80
156
157  Val Lys Arg Ala Ile Arg Ile Leu Thr Ala Arg Lys Arg Leu Val His
158               85              90              95
159
160  Trp Lys Ser Ala Ser Asp Ser Gln Thr Glu Arg Lys Lys Ser Asn His
161              100             105             110
162
163  Met Lys Ser Lys Val Glu Asn Pro Asn Ser Thr Ser Val Arg Ser Ala
164              115             120             125
165
166  Thr Leu Gly His Pro Arg Met Val Met Met Pro Arg Lys Thr Asn Asn
167      130             135             140
168
169
170  (2) INFORMATION FOR SEQ ID NO:3:
171
172      (i) SEQUENCE CHARACTERISTICS:
173          (A) LENGTH: 1012 base pairs
174          (B) TYPE: nucleic acid
175          (C) STRANDEDNESS: single
176          (D) TOPOLOGY: linear
177
178     (ii) MOLECULE TYPE: cDNA
179
180
181     (ix) FEATURE:
182          (A) NAME/KEY: CDS
183          (B) LOCATION: 117..566
184
185     (ix) FEATURE:
186          (A) NAME/KEY: mat_peptide
187          (B) LOCATION: 186..566
188
189
190     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
191
192  TCGACCCACG CGTCCGCTTG GCCTACAGCC CGGCGGGCAT CAGCTCCCTT GACCCAGTGG          60
193
194  ATATCGGTGG CCCCGTTATT CGTCCAGGTG CCCAGGGAGG AGGACCCGCC TGCAGC           116
195
196  ATG AAC CTG TGG CTC CTG GCC TGC CTG GTG GCC GGC TTC CTG GGA GCC        164
197  Met Asn Leu Trp Leu Leu Ala Cys Leu Val Ala Gly Phe Leu Gly Ala
198  -23          -20             -15             -10
199
200  TGG GCC CCC GCT GTC CAC ACC CAA GGT GTC TTT GAG GAC TGC TGC CTG        212
201  Trp Ala Pro Ala Val His Thr Gln Gly Val Phe Glu Asp Cys Cys Leu
202          -5              1               5
203
204  GCC TAC CAC TAC CCC ATT GGG TGG GCT GTG CTC CGG CGC GCC TGG ACT        260
205  Ala Tyr His Tyr Pro Ile Gly Trp Ala Val Leu Arg Arg Ala Trp Thr
```

INPUT SET: S24076.raw

```
206      10                      15                      20                      25
207
208   TAC CGG ATC CAG GAG GTG AGC GGG AGC TGC AAT CTG CCT GCT GCG ATA        308
209   Tyr Arg Ile Gln Glu Val Ser Gly Ser Cys Asn Leu Pro Ala Ala Ile
210                           30                      35              40
211
212   TTC TAC CTC CCC AAG AGA CAC AGG AAG GTG TGT GGG AAC CCC AAA AGC        356
213   Phe Tyr Leu Pro Lys Arg His Arg Lys Val Cys Gly Asn Pro Lys Ser
214                   45                      50                      55
215
216   AGG GAG GTG CAG AGA GCC ATG AAG CTC CTG GAT GCT CGA AAT AAG GTT        404
217   Arg Glu Val Gln Arg Ala Met Lys Leu Leu Asp Ala Arg Asn Lys Val
218               60                      65                      70
219
220   TTT GCA AAG CTC CAC CAC AAC ATG CAG ACC TTC CAA GCA GGC CCT CAT        452
221   Phe Ala Lys Leu His His Asn Met Gln Thr Phe Gln Ala Gly Pro His
222           75                      80                      85
223
224   GCT GTA AAG AAG TTG AGT TCT GGA AAC TCC AAG TTA TCA TCA TCC AAG        500
225   Ala Val Lys Lys Leu Ser Ser Gly Asn Ser Lys Leu Ser Ser Ser Lys
226   90                      95                      100                     105
227
228   TTT AGC AAT CCC ATC AGC AGC AGC AAG AGG AAT GTC TCC CTC CTG ATA        548
229   Phe Ser Asn Pro Ile Ser Ser Ser Lys Arg Asn Val Ser Leu Leu Ile
230                   110                     115                     120
231
232   TCA GCT AAT TCA GGA CTG TGAGCCGGCT CATTTCTGGG CTCCATCGGC              596
233   Ser Ala Asn Ser Gly Leu
234                   125
235
236   ACAGGAGGGG CCGGATCTTT CTCCGATAAA ACCGTCGCCC TACAGACCCA GCTGTCCCCA      656
237
238   CGCCTCTGTC TTTTGGGTCA AGTCTTAATC CCTGCACCTG AGTTGGTCCT CCCTCTGCAC      716
239
240   CCCCACCACC TCCTGCCCGT CTGGCAACTG GAAAGAAGGA GTTGGCCTGA TTTTAACCTT      776
241
242   TTGCCGCTCC GGGGAACAGC ACAATCCTGG GCAGCCAGTG GCTCTTGTAG AGAAAACTTA      836
243
244   GGATACCTCT CTCACTTTCT GTTTCTTGCC GTCCACCCCG GGCCATGCCA GTGTGTCCTC      896
245
246   TGGGTCCCCT CCAAAAATCT GGTCATTCAA GGATCCCCTC CCAAGGCTAT GCTTTTCTAT      956
247
248   AACTTTTAAA TAAACCTTGG GGGGTGAATG GAATAAAAAA AAAAAAAAAA AAAAAA          1012
249
250
251   (2) INFORMATION FOR SEQ ID NO:4:
252
253          (i) SEQUENCE CHARACTERISTICS:
254                 (A) LENGTH: 150 amino acids
255                 (B) TYPE: amino acid
256                 (D) TOPOLOGY: linear
257
258          (ii) MOLECULE TYPE: protein
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/887,977*

INPUT SET: S24076.raw

Line	Error	Original Text